Scientific Symposium on Exposures to
Environmental Contaminants Affecting Children

"New Tools for Environmental Health Research in the Genomics Era"
Plenary Session 2

LCRA - McKinney Roughs
Cedar Creek, Texas

Saturday, October 28, 2000

Samuel Wilson, M.D.
Children's Environmental Health

P R O C E E D I N G S

MS. FIELDS: It is my pleasure to introduce Dr. Samuel Wilson. After I read one of his articles, I went to the Institutes of Medicine/National Academy of Sciences Roundtable on Environmental Health in June. The article had his picture in it, so I sat behind him.

I was able to tap him on the shoulder and talk to him about presenting at this symposium. I do appreciate his responsiveness and being receptive to attending this meeting.

It's my privilege to introduce him.

Since 1996 Dr. Wilson has served as Deputy Director, National Institute of Environmental Health Sciences, National Institute of Health; Chief, DNA Repair and Nucleic Acid Enzymology Section, Laboratory of Structural Biology, NIEHS, NIH.

Prior to taking his current position, Dr. Wilson served as the Founding Director of the Sealy Center for Molecular Science at the University of Texas Medical Branch in Galveston.

He has an M.D. from Harvard Medical School. Dr. Wilson is the author

and co-author of numerous articles and his interests include environmental toxicology, particularly cellular biology, and the effects of the environment on the genome.

Please join me in welcoming Dr. Wilson.

DR. WILSON: Thank you, Ms. Fields.

Well, that was courageous of you reading through that long list of laboratory activities. It's also courageous of you folks in the audience to stick around on such a nice day, to the bitter end. Woodie Kessel reminded me a few minutes ago to make it easy on you.

But the trouble with that very good idea, Woodie, is that I had already pulled my slides back at the hotel before coming here.

So, I'm afraid you are pretty much stuck this afternoon with some detailed comments about genes and genetics and gene-environment interactions. This is the topic I want to focus on this afternoon.

I want to couch my discussion of this topic of gene-environment interactions in the context of our Nation's environmental health research needs. So, before I launch into my slides, let me first make a few general comments about environmental health.

I will start by saying it straight out: We know that the environment is making children sick. Yet, in many cases the specific mechanisms are still not clear. We have information gaps. This lack of information hinders us as a Nation in addressing the public health burden secondary to environmental exposures.

Why does the information gap hinder us? There is the reason Woodie Kessel called to our attention during the panel discussion, and other folks called to

our attention during the question-and-answer period this morning: To develop effective policies for prevention of disease, we use a standard in this country of knowing cause-and-effect relationships, knowing mechanisms, and being able to justify the evolution of policies that can prevent disease.

Environmental factors are making us sick. I take this as a matter of faith, if you will, and believe that as we accomplish more research in the future, we will find out more about the mechanisms and perceive cause-and-effect relationships. Now, why am I so confident in making this statement about the role of the environment?

The notion is simple: chronic disease arises from a combination of genetic factors and environmental factors. We know that only a small portion of overall disease burden is attributable to the genetic diseases, the strongly penetrant inherited traits. That is, diseases for which an inherited gene is strongly associated with disease development at a predictable time in life.

In the case of the inherited diseases, the genetic factor is predominant. Environmental factors can modify the time of penetrance and certain other details, but we can know the disease will appear and with a more or less predictable outcome.

As I said, the inherited diseases represent a relatively small percentage, so far as overall burden of all disease is concerned. Most of the disease burden we face, from the chronic diseases, occurs from

environmental exposures in combination with our individual genetic susceptibility.

Historically, infectious agents, of course, were a major factor in overall disease burden in the U.S. And, infectious disease is still a factor today. But not in the same way percentage wise as it was, say, 200 years ago. As we learn more and more about disease mechanism and ways to understand and quantify harmful

exposures, the importance of gene-environment interaction will be considered paramount throughout the fields of medicine and medical research.

We need to fill in the information gaps: The information gaps in terms of what our children are exposed to; what approaches should we use to maintain accurate "relational" databases: Databases we can use at the community level to find out answers to such questions as: "What are our kids exposed to?" and "What methods are available for rapid assessment of exposures?" Today, when a child goes into the doctor's office to get a serum measurement for lead or some other toxicant it, unfortunately, is not a very happy time because the test will hurt. It's not a happy time even for the doctor because the doctor doesn't want to conduct a painful test that may not be clearly indicated. We need new assays for environmental exposures; new methods so that children can be quickly examined, without pain. It shouldn't always be necessary to go to the doctor's office for these assays; they should be done in the home or in the school or in a field station. And, again, it should actually be fun, to get a test to find out the status of our environmental exposures. Now, with only a few exceptions, we don't have such assays at the present time. I am talking about a vision for the future. We need to consider new research toward developing non-invasive techniques so as to be able to measure exposures. Clearly, since we have the engineering to develop the internet and to send rockets to Mars and take photographs, there's no doubt that we can develop ways to measure the priority environmental agents that are causing disease. I would hope the field of engineering and the field of biomedical research can collaborate to address this need in the near future, to develop fundamentally more powerful assays for exposure assessment. We also need to build partnerships with industry to do this, because the traditional

ways of doing business in biomedical research, basically leaving industry out, won't work as well.

In summary, medical researchers need to partner with industry. We need to partner with the physical sciences, engineering sciences, and chemical sciences to find new exposure assessment approaches.

A challenge we have heard here at this meeting, and one we constantly hear in the environmental health field, is how do we understand the exposure-disease relationship? There again, to answer the question, we need exposure assays, and new ways of understanding the balance between exposure to a toxicant and the cellular response. We need new molecular endpoints for describing what diseases actually are, in molecular terms.

I will now turn to one approach toward addressing this challenge of developing new assays. It is fair to say that we are now in a revolution of new technology in the area of genome research.

I will discuss applications of genome research in the environmental health sciences. The NIEHS has initiated two new programs in this area. These programs have an opportunity to arm us with a completely new set of assays and ways of understanding environmental disease. Obviously, the word “genomics” is here to stay, and we are in the midst of the “genomics revolution.” In the biomedical sciences, the new sequence information on the human genome represents extraordinary opportunity for a whole range of applications. One of these applications is the research initiative at NIEHS called the Environmental Genome Project. With this project, we are stimulating research on better understanding of gene-environment

interactions. Genetic variations are being identified that cause humans to be unusually susceptible to an environmental exposure.

Some of the scientific underpinning for the Environmental Genome Project comes from the field of pharmacology, where we know that different individuals are more susceptible to certain drugs than other individuals. The underpinning also comes from the field of cancer risk assessment, where we know that some individuals have a higher cancer risk to a certain type of exposure than other individuals. Genomic polymorphisms, a term we use for DNA variations, in metabolism genes are common in the population and vary by ethnic group. These metabolism gene variations account for a lot of the difference between individuals in drug response and in cancer risk. Actually, a number of gene categories other than metabolism genes have also been identified as being responsible for individual differences in susceptibility. This background information clearly tells us that an individual's genetic composition in combination with exposure, such as to a drug, is very important in determining one's individual health. Some features of the Environmental Genome Project are summarized on this slide (Slide 75646). The project focuses in the field of epidemiology, as we seek to develop sharper, more precise tools for epidemiologists to use (in their population-based studies). Improving technologies, such as DNA-based technologies, and finding new assays for early disease is the point of this first bullet. The second bullet, optimizing study design, is a very challenging point. We will need large population cohorts or groups with availability to both DNA samples and medical records. Finding ways to understand these challenges and acquire the requisite research resources is the point of these 2nd and 3rd bullets. And finally, this last point, addressing the ethical,

legal, and social implications of the new genetic research. This challenge will be a feature of the genomic age and is a challenge we must deal with, as we move ahead in the environmental health field and in other fields of biomedical research. These are just some of the salient features of the Environmental Genome Project.

Now, for the second of the new NIEHS genomics initiatives: The initiative is called “toxicogenomics.” The initiative has not yet been announced publicly. It's still under development. It is a program that will allow researchers to make use of the new sequence of the human genome in the field of toxicology and environmental health.

We plan to use an instrument, termed the “National Center for Toxicogenomics” to facilitate and stimulate research in this area. This Center will be managed and housed on our campus in Research Triangle Park. Let me first explain the definition of toxicogenomics (Slide A). It's a scientific field that studies how the entire genome responds to environmental stress or environmental toxicants. Toxicogenomics combines studies of genetics, as in the Environmental Genome Project, with genomic scale messenger RNA expression, cell-wide and tissue-wide protein expression, and bioinformatics to understand the role of gene-environment interactions in disease. Thus, toxicogenomics will be a field in environmental health sciences focusing on genomic scale expression of messenger RNA and proteins and informatics.

The research area is new because, for the first time, we have access to genomic-scale DNA sequence information for humans and many model organisms. One way to reduce this new information to practice in health research is to conduct programs like the toxicogenomics initiative.

Now, let's look at the list in the next slide to see what the initiative is all about (Slide B). The proposed goal is to facilitate genome-wide expression studies and gene-expression studies of gene-environment interactions in disease etiology.

We propose to do this primarily through the university/academic community, as with other NIH programs. Other features of the initiative are listed: To improve technologies for genome-wide analysis of expression of messenger RNA and protein; optimizing study design and understanding the most important problems to tackle. In other words, we must phrase research questions in this initiative that are the most exciting questions possible and this will be a very important challenge. Finally, developing a national research capacity in this new area. And the last point, addressing the application of toxicogenomics so as to inform public policy, in order to achieve disease prevention and intervention.

Now, on the next slide (Slide C), we see a model of one of the anticipated outcomes of this research: a database where information from toxicogenomics research will be made widely available, both to the scientific community and practicing physicians, and also to the public and policy makers. We envision a database that will be maintained continuously. In other words, with timely curation for accuracy and for searchability, of exposure information and disease outcome. There are five major components or sectors of information for this database on the exposure/disease paradigm. Molecular indicators of exposure; for example, lead exposure is not only a measurement of body burden of lead, but would be a set of molecular markers indicating that lead exposure has occurred. Other information components or sectors are genetic variation, messenger RNA expression patterns and protein expression patterns, and molecular indicators of toxicity or early

disease. These five sectors of information will populate the database. The database will be relational from the standpoint of allowing one to search from a disease endpoint, or for an exposure endpoint or for a genetic variation endpoint, and even for a messenger RNA profile difference endpoint.

Next, I will discuss four case-studies illustrating that genomic scale messenger RNA expression can actually work. In the first, data were taken from a paper published recently in Nature that was part of a group of three or four papers that indicated to the cancer biology field, and in the environmental health sciences field, that genomic-scale messenger RNA profiling should be taken seriously (Slide D). The aim of the study, which was done by a collaborative team at the National Cancer Institute and Stanford University, was to examine gene expression for a type of lymphoma in a group of patients. Expression of about 10,000 different genes was measured simultaneously, using DNA chip technology for hybridization of messenger RNAs, to measure expression levels for all of these genes. At the end of the measurement phase of the experiment, they (and a group at NIEHS, Li *et al.* at NIEHS) analyzed the results and boiled the information down to a subset of 50 different genes. The 50 genes selected allowed the investigators to distinguish between activated and germinal center diffuse large β-cell lymphoma, two types of lymphoma cells. The activated cell pattern for these 50 genes is shown in the left hand part of the slide. The germinal center cell pattern for the 50 genes is over here on the right hand part of the slide. You can see just by glancing at this slide that the two patterns are quite different. Thus, this technology appeared to be powerful enough to allow a pathologist/oncologist to distinguish between these types of β-cell lymphoma.

Furthermore, another interesting point emerged in this paper. In the middle zone of the slide, there appears to be a transition from one cell pattern to the other cell pattern. These samples corresponding to the middle zone are candidate cells to be in transition from one state to the other. As I understand it, it is not possible with the existing cytological markers for β-cell lymphoma to diagnose this intermediate stage. Right away, these experiments suggested a new experimental capacity for the field of oncology/cancer biology, making use of genome-wide messenger RNA expression. The data analysis shown in this slide, by the way, was done by a group at NIEHS, shown here at the bottom.

This next slide (Slide E) illustrates another study, one published by Arnold Levine and his colleagues at Rockefeller University. They obtained colon cancer specimens and specimens of adjacent normal mucosa and subjected them to messenger RNA profiling. They compared the cancer tissue patterns in this part of the slide, with adjacent normal mucosal tissue in the other part of the slide.

One can easily see that normal tissue gives a different pattern than that obtained with colon cancer tissue. Furthermore, there aren't false positives or false negatives in this entire set of specimen patterns. The patient samples or specimens are listed across the top. The 50 genes that were selected to diagnose the difference between the normal and cancer samples are listed here on the side. So, the test appeared to be absolutely diagnostic, and very powerful in terms of characterizing normal colonic mucosa and colon cancer.

Finally, here is another slide on the same theme. Except this is a case study about leukemia. It was published in Science less than a year ago by Eric Lander and his colleagues at MIT.

Once again, the analysis was done with 50 distinguishing genes listed on the right-hand side. Across the top is the collection of samples from different patients (Slide F). These investigators examined two different types of leukemia. One called “AML” and the other “ALL.” In the left-hand part of the slide one can see that specimens from patients corresponding to AML are clustered, whereas here in another part of the slide the patient’s specimens corresponding to ALL are clustered. Thus, these two different types of leukemia could be distinguished. Furthermore, the ALL cases could be sub-categorized based upon the messenger RNA expression patterns. These case studies demonstrate the point, in my view, that genomic-scale messenger RNA expression is a powerful tool for cancer biology.

But what about use of this approach in toxicology and environmental health sciences (Slide G)? One idea is to expose cells or animals to an unknown or suspected toxicant and then determine the expression patterns. One can then ask, “Do the patterns match with those produced by known toxicants?” We can see in this diagram, prepared by Cindy Afshari and Rick Paules at NIEHS, that there is a match with these red dots across the top, which would suggest in this experiment that the unknown toxicant is an oxidative stress-type toxicant. Basically, this is the idea of how we hope to begin to use messenger RNA chip technology in the field of toxicology. The next slide (Slide H) shows an example of a case study on reduction to practice by this same NIEHS group, working in the intramural laboratories. They examined the effect of four different toxic compounds on messenger RNA expression in rat liver. Animals were exposed to one of the four chemicals and then tested. The question was, “Would the messenger RNA expression analysis reveal different patterns for four different toxic chemicals?” Here on the left-hand side, you can see that the first

chemical gives a characteristic pattern. On the right-hand side, a different chemical, phenobarbital, gives a completely different pattern, for rat liver gene expression. So clearly, this technology seems to be powerful enough to be used in toxicology. We expect to be able to use this kind of approach in further understanding of the gene expression specificity for exposures to a whole variety of environmental toxicants.

Moving on, how will this new toxicogenomics initiative work? As I said, the research will be conducted mainly at universities across the country. This slide shows (Slide I) our model, at the present time, for how the National Center for Toxicogenomics will function once it's up and running. We envision a research consortium and a “central contractor” to collate toxicogenomics data and to assist with cross-laboratory standardization for the data. Centers (“Academic Research Members”) in the extramural community that will do hybridizations and submit the data to this central contractor. We will also incorporate other toxicology information from the various NCT centers, other NIEHS and NIH centers and from scientists working under the support of RO-1 grants and PPG grants. We also wish to incorporate data from industry, as industry is already very active in the use of chip technology for messenger RNA profiling.

Now, one feature of the type of consortium envisioned for the NCT is the opportunity to know that assays in one center can be compared with the same assay done in another center. Thus, across-platform standardization and validation of assays is a fundamental challenge, in order to build a reliable database in this field.

In summary, the use of messenger RNA profiling and similar profiling for protein expression offers a fundamental new revolution in the environmental health sciences. We hope to define environmental disease in molecular terms and define

exposures in molecular terms. The power and precision of this approach will be vastly greater than our current tools. I think our research opportunity in this field in the future will be essentially unlimited.

In closing, let me say that it's a tremendously exciting time to be working in environmental health sciences.

The name of the game in this field is disease prevention. And, at the same time, the name of the game is gene-environment interactions. We need to go on the “warpath” as a scientific field, as groups of concerned physicians and concerned citizens to get this work done: to do more research and to get these toxicogenomic databases constructed. So with that, I'll close. Thank you very much Janie for this opportunity; congratulations to you and your colleagues on this fine symposium and on the Children’s Environmental Health Institute.